

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

SCIENCE @ DIRECT®

J. Math. Anal. Appl. 302 (2005) 402–416

---

*Journal of*  
 MATHEMATICAL  
 ANALYSIS AND  
 APPLICATIONS
 

---

[www.elsevier.com/locate/jmaa](http://www.elsevier.com/locate/jmaa)

# The absolute convergence of left product representations for $\Phi' = \mathbf{A}(t)\Phi$

Igor Najfeld\*, William Lakin

*Department of Mathematics and Statistics, University of Vermont, Burlington, VT, USA*

Received 2 October 2003

Submitted by J. Henderson

---

## Abstract

It was shown in [Quart. Appl. Math. LXI (2003) 537–564] that the fundamental matrix of solutions for a system  $\Phi' = \mathbf{A}(t)\Phi$  possesses three left ordered product representations. Here, for each product, we derive a sufficient condition for its absolute convergence by first constructing a scalar fixed point problem and then computing its positive solution.

© 2004 Elsevier Inc. All rights reserved.

*Keywords:* Left product representation; L-Derivative; Absolute convergence of matrix products

---

## 1. Introduction

The fundamental matrix of solutions,  $\Phi(t)$ , of a  $d$ -dimensional system of linear differential equations,

$$\Phi' = \mathbf{A}(t)\Phi, \quad \Phi(0) = \mathbf{I}, \quad (1.1)$$

with  $\mathbf{A}(t), \Phi(t) \in \mathbb{C}^{d \times d}$ ,  $\mathbf{I}$  = identity matrix, was shown in [3] to have three product representations. The goal of this paper is to derive sufficient conditions for the absolute convergence of each product.

---

\* Corresponding author.

*E-mail addresses:* [igor@emba.uvm.edu](mailto:igor@emba.uvm.edu) (I. Najfeld), [lakin@emba.uvm.edu](mailto:lakin@emba.uvm.edu) (W. Lakin).

Briefly, if  $\Phi_0(t)$  is an initial approximation to the solution of (1.1) and the classical relative error function is written as  $\mathbf{R}(t) = \mathbf{I} - \Phi_0(t)\Phi^{-1}(t)$ , then the associated left ordered product representation for  $\Phi(t)$  is given by

$$\Phi(t) = \left( \prod_{k=0}^{\infty} (\mathbf{I} - \mathbf{R}_k(t))^{-1} \right) \Phi_0(t). \quad (1.2)$$

The matrix sequence  $\{\mathbf{R}_k\}_{k \geq 0}$  is generated by a single step iterative process (see (2.7)) and ‘left ordered’ means that the  $(k+1)$ st factor is located to the left of the  $k$ th factor.

Similarly, the modified relative error, defined by  $\mathbf{R}(t) = \Phi(t)\Phi_0^{-1}(t) - \mathbf{I}$ , generates its own left ordered product representation for  $\Phi(t)$  in the form

$$\Phi(t) = \left( \prod_{k=0}^{\infty} \mathbf{I} + \mathbf{R}_k(t) \right) \Phi_0(t). \quad (1.3)$$

When  $\Phi(t)$  is unitary or symplectic, as is the case in a number of applications, neither the representation (1.2) nor (1.3) is able to preserve these properties in a truncated product. It turns out that the A-relative error, defined by  $\mathbf{R}(t) = 2(\Phi(t) - \Phi_0(t))(\Phi(t) + \Phi_0(t))^{-1}$ , gives rise to the left ordered product representation,

$$\Phi(t) = \left( \prod_{k=0}^{\infty} \frac{2 + \mathbf{R}_k(t)}{2 - \mathbf{R}_k(t)} \right) \Phi_0(t), \quad (1.4)$$

which does preserve both properties in any truncated product, provided the initial approximation  $\Phi_0$  possesses it [3]. Here and in what follows, the numerator and the denominator matrices are written in the rational form when they commute, otherwise the inverse notation is used. Whenever it is clear what the factors are, the parentheses within a product are omitted, as in (1.3). Bold **2** stands for  $2\mathbf{I}$ , as in (1.4). Except for  $\mathbf{R}_0$ , the sequences  $\{\mathbf{R}_k\}_{k \geq 0}$  are different for each product, but the basic structure of all three iterations, (2.7), (3.2), (4.5), that generates them is fairly similar.

The notation used here is the same as in [3], but some definitions will be restated to maintain continuity. An initial approximation,  $\Phi_0(t)$ , is said to be *admissible* if it is bounded, invertible and differentiable for all  $t$  in some finite region  $D$  and satisfies  $\Phi_0(0) = \mathbf{I}$ . The L-derivative of  $\Phi_0(t)$  is then defined by  $\mathbf{L}[\Phi_0(t)] = \Phi_0'(t)\Phi_0^{-1}(t) = \mathbf{A}_0(t)$ . The difference between the two L-derivatives,  $\mathbf{A} = \mathbf{L}[\Phi]$  and  $\mathbf{A}_0 = \mathbf{L}[\Phi_0]$ , is termed the L-perturbation and is denoted by  $\Delta_0(t) = \mathbf{A}(t) - \mathbf{A}_0(t)$ . It is assumed that  $\Delta_0(t)$  is a known function and that near the origin the norm inequality,

$$\|\Delta_0(t)\| \leq \delta t^p, \quad p \geq 0, \quad (1.5)$$

holds for some integer  $p$ .

By the admissibility assumptions on  $\Phi_0(t)$ , the function

$$\kappa_0(t) = \max_{0 < \sigma \leq t} \|\Phi_0(t, \sigma)\| \|\Phi_0^{-1}(t, \sigma)\| \quad (1.6)$$

exists for all  $t \in D$ . Its upper bound over  $t \in D$  is known (computable), hence there is a constant  $K \geq 1$  such that  $\kappa_0(t) \leq K$ . With this notation, one may define the function,

$$\omega(t) = K \delta \frac{t^{p+1}}{p+1}, \quad (1.7)$$

which will often be used as an independent variable.

All the infinite series appearing here can be bounded by the function,

$$f(z) = g(z) + g(z^2) + g(z^4) + \cdots = \sum_{k=0}^{\infty} g(z^{2^k}), \quad |z| < 1. \quad (1.8)$$

The generator function,  $g$ , is analytic within the unit circle and satisfies the condition  $g(0) = 0$ . The simplest such case,  $g(z) = z$ , is cited in [2], where it is shown that the singularities of  $f(z)$  are dense on  $|z| = 1$ , making it the natural boundary.

A function  $f$  in (1.8), having no other singularities within its natural boundary, satisfies infinitely many functional equations of the form,

$$f(z) = g(z) + f(z^2) = g(z) + g(z^2) + f(z^4) = \cdots = \sum_{k=0}^{m-1} g(z^{2^k}) + f(z^{2^m}).$$

Thus,  $f(z^{2^m})$  is the exact remainder for the finite expansion of length  $m$ . Since  $g(0) = 0$ , then, near the origin,  $f(z^{2^m}) = O(g(z^{2^m})) = O(z^{2^m})$ , or higher. Consequently, one can determine a rather accurate bound on the remainder  $f(z^{2^m})$  for a small  $m$ . In all the cases considered here,  $g$  is a simple rational function which means that the evaluation of  $f(z)$  is very rapid for any  $z$  within its natural boundary, except possibly for the points arbitrarily close to it. In particular, each proof of the absolute convergence will be shown to depend on a specific function  $f(r)$ , with real  $r \in [0, 1)$ .

## 2. The classical relative error and its product representation

Let  $\Phi_0(t)$  be an admissible initial approximation,  $\mathbf{A}_0(t)$  its L-derivative, and  $\Delta_0(t)$  its L-perturbation. If the classical relative error matrix is defined by  $\mathbf{R}(t) = \mathbf{I} - \Phi_0(t)\Phi_0^{-1}(t)$ , then

$$\Phi(t) = (\mathbf{I} - \mathbf{R}(t))^{-1} \Phi_0(t). \quad (2.1)$$

The differential equation for  $\mathbf{R}(t)$  can be written in terms of the commutator,  $[\mathbf{A}_0, \mathbf{R}] = \mathbf{A}_0\mathbf{R} - \mathbf{R}\mathbf{A}_0$ , as follows:

$$\mathbf{R}' = \Delta_0 + [\mathbf{A}_0, \mathbf{R}] - \mathbf{R}\Delta_0, \quad \mathbf{R}(0) = \mathbf{0}. \quad (2.2)$$

Its integrable part,  $\mathbf{R}_0(t)$ , satisfies the linear differential equation

$$\mathbf{R}_0' = \Delta_0 + [\mathbf{A}_0, \mathbf{R}_0], \quad \mathbf{R}_0(0) = \mathbf{0}. \quad (2.3)$$

Hence, it is explicitly given by the quadrature of a known matrix function,

$$\mathbf{R}_0(t) = \int_0^t \Phi_0(t, \sigma) \Delta_0(\sigma) \Phi_0^{-1}(t, \sigma) d\sigma. \quad (2.4)$$

Once the first approximation,  $\mathbf{R}_0(t)$ , is computed from (2.4), it may be substituted into (2.1) to correct  $\Phi_0(t)$ ,

$$\Phi_1(t) = (\mathbf{I} - \mathbf{R}_0(t))^{-1} \Phi_0(t). \quad (2.5)$$

The matrix  $\Phi_1(t)$  may be viewed as another initial approximation, provided it is admissible, an important point which will be addressed later. The computation of the matrix  $\Delta_1(t) = \mathbf{A}(t) - \mathbf{L}[\Phi_1(t)]$  shows [3] that it is a linear function of  $\Delta_0(t)$ ,

$$\Delta_1(t) = -(\mathbf{I} - \mathbf{R}_0(t))^{-1} \mathbf{R}_0(t) \Delta_0(t). \quad (2.6)$$

Equations (2.4), (2.5) and (2.6) define the iterative process that refines any admissible initial approximation,  $\Phi_0(t)$ , and its L-perturbation,  $\Delta_0(t)$ :

$$\begin{aligned} \mathbf{R}_k(t) &= \int_0^t \Phi_k(t, \sigma) \Delta_k(\sigma) \Phi_k^{-1}(t, \sigma) d\sigma, \\ \Phi_{k+1}(t) &= (\mathbf{I} - \mathbf{R}_k(t))^{-1} \Phi_k(t), \\ \Delta_{k+1}(t) &= -(\mathbf{I} - \mathbf{R}_k(t))^{-1} \mathbf{R}_k(t) \Delta_k(t). \end{aligned} \quad (2.7)$$

The second equation in (2.7) generates left ordered infinite product representation for the fundamental matrix of solutions,

$$\Phi(t) = \lim_{m \rightarrow \infty} \left( \prod_{k=0}^m (\mathbf{I} - \mathbf{R}_k(t))^{-1} \right) \Phi_0(t). \quad (2.8)$$

The convergence of this product depends on the function  $\omega(t)$  in (1.7).

**Theorem 1.** *The infinite product (2.8) is absolutely convergent provided*

$$\omega(t) < \omega_* = 0.79681213002\dots \quad (2.9)$$

**Proof.** By rewriting each factor of the product (2.8) in the standard form,  $(\mathbf{I} - \mathbf{R}_k(t))^{-1} = \mathbf{I} + (\mathbf{I} - \mathbf{R}_k(t))^{-1} \mathbf{R}_k(t)$ , it follows from the theorem in [1, Chapter 8] that the product (2.8) converges absolutely provided the associated matrix series,  $\sum_{k=0}^{\infty} (\mathbf{I} - \mathbf{R}_k(t))^{-1} \mathbf{R}_k(t)$ , converges absolutely. Upon taking the norm of this series, and then using the inequality  $\|(\mathbf{I} - \mathbf{R}_k(t))^{-1}\| \leq (1 - \|\mathbf{R}_k(t)\|)^{-1}$ , as long as  $\|\mathbf{R}_k(t)\| < 1$ , this condition becomes

$$\sum_{k=0}^{\infty} \|(\mathbf{I} - \mathbf{R}_k(t))^{-1} \mathbf{R}_k(t)\| \leq \sum_{k=0}^{\infty} \frac{\|\mathbf{R}_k(t)\|}{1 - \|\mathbf{R}_k(t)\|} < \infty. \quad (2.10)$$

*Step 1.* We start with the assumption that there is a positive constant  $K$  (the same as in (1.7)) such that

$$\kappa_k(t) = \max_{0 < \sigma \leq t} \|\Phi_k(t, \sigma)\| \|\Phi_k^{-1}(t, \sigma)\| \leq K, \quad k \geq 0. \quad (2.11)$$

The justification for this assumption is deferred until the fifth step where the constant  $K$  is computed provided the condition (2.9) holds.

*Step 2.* To set up an induction, take the norm of the first equation in (2.7), then use the bounds (1.5) and (2.11), followed by the explicit integration. The resulting bound on  $\|\mathbf{R}_0(t)\|$  is

$$\begin{aligned}
r_0(t) &= \|\mathbf{R}_0(t)\| \leq \int_0^t \|\Phi_0(t, \sigma)\| \|\Delta_0(\sigma)\| \|\Phi_0^{-1}(t, \sigma)\| d\sigma \\
&\leq K \int_0^t \delta \sigma^p d\sigma = \frac{K \delta t^{p+1}}{p+1} = \omega(t).
\end{aligned} \tag{2.12}$$

Next, substitute the inequality (2.12) and the expression  $\omega'(t)/K = \delta t^p$  into the norm of the third equation in (2.7) to obtain the bound on  $\|\Delta_1(t)\|$ ,

$$\begin{aligned}
\delta_1(t) &= \|\Delta_1(t)\| \leq \|(\mathbf{I} - \mathbf{R}_0(t))^{-1}\| \|\mathbf{R}_0(t)\| \|\Delta_0(t)\| \\
&\leq \frac{r_0(t)}{1 - r_0(t)} \delta t^p \leq \frac{\omega(t)}{1 - \omega(t)} \frac{\omega'(t)}{K}.
\end{aligned} \tag{2.13}$$

In the second iterative step, use the inequalities (2.11) and (2.13) to calculate the bound on  $r_1(t) = \|\mathbf{R}_1(t)\|$  as follows:

$$\begin{aligned}
r_1(t) &\leq \int_0^t \|\Phi_1(t, \sigma)\| \|\Delta_1(\sigma)\| \|\Phi_1^{-1}(t, \sigma)\| d\sigma \\
&\leq \int_0^t \frac{\omega(\sigma)}{1 - \omega(\sigma)} d(\omega(\sigma)) = \phi(\omega(t)) - \phi(\omega(0)).
\end{aligned} \tag{2.14}$$

Here, the last integral is evaluated exactly in terms of the analytic function

$$\phi(\omega) = -\omega - \ln(1 - \omega) = \sum_{j=2}^{\infty} \frac{\omega^j}{j}, \quad |\omega| < 1. \tag{2.15}$$

By (1.7),  $\omega(0) = 0$ , hence  $\phi(\omega(0)) = \phi(0) = 0$ , and (2.14) is rewritten as

$$r_1(t) \leq \phi(\omega(t)) = \phi^{[1]}(\omega(t)). \tag{2.16}$$

Setting  $r_0(t) \leq \omega(t) = \phi^{[0]}(\omega(t))$  indicates that the inequality (2.16) could be the first step in an induction to prove a general bound,

$$r_k(t) \leq \phi^{[k]}(\omega(t)), \quad \phi^{[k]}(\omega) = \phi(\phi(\dots\phi(\omega)\dots)). \tag{2.17}$$

Here,  $\phi^{[k]}(\omega)$  denotes  $k$ th functional iterate of  $\phi$  for any positive integer  $k > 0$ . To show this, assume that in the  $k$ th step  $r_k(t)$  is already bounded, as in (2.17), and that, by analogy to (2.13), the following inequality holds:

$$\delta_k(t) \leq \frac{1}{K} \frac{d\phi^{[k]}(\omega(t))}{dt}. \tag{2.18}$$

The recurrence relation for  $(d/dt)\phi^{[k]}(\omega(t))$  is obtained from the definition,  $\phi^{[k]}(\omega) = \phi(\phi^{[k-1]}(\omega))$ , and from  $d\phi(\omega)/d\omega = \omega/(1 - \omega)$ , which follows from (2.15). Therefore, for each  $k > 0$ ,

$$\frac{d\phi^{[k]}(\omega(t))}{dt} = \frac{d\phi(z)}{dz} \frac{dz}{dt} = \frac{z}{1 - z} \frac{dz}{dt}, \quad z(t) = \phi^{[k-1]}(\omega(t)). \tag{2.19}$$

Then the substitution of the bounds (2.17), (2.18), and the (2.17), (2.19), into the expression for the bound on  $\delta_{k+1}(t) = \|\mathbf{\Delta}_{k+1}(t)\|$  yields

$$\begin{aligned}\delta_{k+1}(t) &\leq \|(\mathbf{I} - \mathbf{R}_k(t))^{-1}\| \|\mathbf{R}_k(t)\| \|\mathbf{\Delta}_k(t)\| \\ &\leq \frac{\phi^{[k]}(\omega(t))}{1 - \phi^{[k]}(\omega(t))} \frac{1}{K} \frac{d\phi^{[k]}(\omega(t))}{dt} = \frac{1}{K} \frac{d\phi^{[k+1]}(\omega(t))}{dt}.\end{aligned}\quad (2.20)$$

This shows that the inequality (2.18) holds for all  $k > 0$ . The substitution of the bound (2.20) into the expression for  $r_{k+1}(t) = \|\mathbf{R}_{k+1}(t)\|$  results in

$$\begin{aligned}r_{k+1}(t) &\leq \int_0^t \|\Phi_{k+1}(t, \sigma)\| \|\mathbf{\Delta}_{k+1}(\sigma)\| \|\Phi_{k+1}^{-1}(t, \sigma)\| d\sigma \\ &\leq \int_0^t \frac{d\phi^{[k+1]}(\omega(\sigma))}{d\sigma} d\sigma = \phi^{[k+1]}(\omega(t)).\end{aligned}$$

The last equality follows from the fact that  $\phi^{[k]}(0) = 0$  for all  $k > 1$ , whenever  $\phi^{[1]}(0) = 0$ . This proves that the inequality (2.17) holds for all  $k > 0$ .

The positivity of the coefficients in the series (2.15) makes it clear that  $\phi$  is a real, positive and monotonically increasing function for  $\omega \in (0, 1)$ . It is less clear that the functions  $\phi^{[k]}$ , for  $k > 1$ , have the same properties over the same interval. Already for  $k = 2$ , it is apparent from

$$\phi^{[2]}(\omega) = -\phi(\omega) - \log(1 - \phi(\omega))$$

that the size of  $\phi(\omega)$  must be restricted to ensure the existence and nonnegativity of  $\phi^{[2]}(\omega)$ . This means that  $\phi(\omega) < 1$ . Solving this inequality for  $\omega$  generates the first restriction:  $\omega < \omega_1 \approx 0.84140566$ . Continuing in this way to the  $(k + 1)$ st step and the function

$$\phi^{[k+1]}(\omega) = -\phi^{[k]}(\omega) - \log(1 - \phi^{[k]}(\omega)),$$

the existence and nonnegativity of  $\phi^{[k+1]}(\omega)$  requires that  $\phi^{[k]}(\omega) < 1$ . Since  $\phi^{[k]}(\omega) = \phi^{[k-1]}(\phi(\omega)) = \dots = \phi(\phi^{[k-1]}(\omega))$ , it follows that the inequality  $\phi^{[k]}(\omega) < 1$  implies the inequality  $\phi(\omega) < \omega_{k-1}$ . The positive solution to this last nonlinear inequality, denoted by  $\omega_k$ , is the  $k$ th restriction on  $\omega$ , or the next entry in the monotonically decreasing sequence,  $\{\omega_k\}_{k=1}^\infty$ . This sequence is equivalent to a (slightly altered) standard fixed point iteration,

$$\phi(\omega_{k+1}) = \omega_k, \quad \omega_0 = 1, \quad (2.21)$$

for the equation

$$\phi(\omega) = \omega. \quad (2.22)$$

The iteration (2.21) converges to  $\omega_* = 0.7968121\dots$ , the unique positive solution of (2.22). The inequality (2.9) is therefore a sufficient condition for the existence and nonnegativity of all functional iterates  $\phi^{[k]}$ .

*Step 3.* The functions  $\phi^{[k]}$  satisfy several useful properties. Over the interval  $[0, \omega_*]$ , the function  $\phi$  can be bounded by the quadratic polynomials

$$\frac{\omega^2}{2} \leq \phi(\omega) \leq c\omega^2. \quad (2.23)$$

The lower bound follows from the series (2.15). The constant  $c$  in the upper bound is computed from the interpolatory condition (2.22),

$$\phi(\omega_*) = c\omega_*^2 = \omega_*, \quad (2.24)$$

hence,

$$c = \omega_*^{-1} = 1.25500097 \dots \quad (2.25)$$

The analysis of the difference function,  $\epsilon(\omega) = c\omega^2 - \phi(\omega)$ , shows that its only zeros are the endpoints of the interval  $[0, \omega_*]$  and its only maximum there occurs at  $1 - \omega_*/2 \approx 0.60159$ , with  $\epsilon_{\max} \approx 0.13551$ . This proves the upper bound in (2.23).

The three properties that hold on the interval  $[0, \omega_*]$ , the monotonicity of  $\phi$ , the inequality (2.23), and  $\phi^{[1]}(\omega) \leq \phi^{[0]}(\omega) = \omega$ , can be shown to extend to function  $\phi^{[2]}$ , which is also monotonically increasing, satisfies  $\omega^4/8 \leq \phi^{[2]}(\omega) \leq (c\omega)^4/c$ , and  $\phi^{[2]}(\omega) \leq \phi^{[1]}(\omega)$ . It follows by induction that, for each  $k > 0$ , the function  $\phi^{[k]}$  is monotonically increasing, bounded by

$$2\left(\frac{\omega}{2}\right)^{2^k} \leq \phi^{[k]}(\omega) \leq \frac{1}{c}(c\omega)^{2^k}, \quad \omega \in [0, \omega_*], \quad (2.26)$$

and the functional iterates satisfy the following set of inequalities:

$$0 \leq \dots \leq \phi^{[k]}(\omega) \leq \dots \leq \phi^{[1]}(\omega) \leq \omega \leq \omega_* < 1. \quad (2.27)$$

*Step 4.* With the help of the inequalities (2.12) and (2.17), the second series in (2.10) may be bounded as follows:

$$\sum_{k=0}^{\infty} \frac{\|\mathbf{R}_k(t)\|}{1 - \|\mathbf{R}_k(t)\|} \leq \sum_{k=0}^{\infty} \frac{r_k(t)}{1 - r_k(t)} \leq \sum_{k=0}^{\infty} \frac{\phi^{[k]}(\omega(t))}{1 - \phi^{[k]}(\omega(t))}. \quad (2.28)$$

In the last sum, all the terms exist since the denominators are bounded away from zero, by (2.27), provided  $\omega \in [0, \omega_*]$ . Therefore, the substitution of the upper bounds (2.26) is valid and it leads to

$$\sum_{k=0}^{\infty} \frac{\phi^{[k]}(\omega)}{1 - \phi^{[k]}(\omega)} \leq \frac{\omega}{1 - \omega} + \sum_{k=1}^{\infty} \frac{(c\omega)^{2^k}}{c - (c\omega)^{2^k}} = f(c\omega). \quad (2.29)$$

Here, the analytic function (see (1.8)),

$$f(z) = \frac{z}{c - z} + \frac{z^2}{c - z^2} + \frac{z^4}{c - z^4} + \dots = \sum_{k=0}^{\infty} \frac{z^{2^k}}{c - z^{2^k}}, \quad |z| < 1,$$

generated by  $g(z) = z/(c - z)$ , is real, nonnegative and monotonically increasing on the real interval  $z \in [0, 1)$ . As  $\omega \rightarrow \omega_*$ ,  $c\omega \rightarrow 1$  which is on the natural boundary of  $f$ . Hence,

as long as  $\omega(t) \in [0, \omega_*)$ , the function  $f(c\omega)$  provides a bound on the series (2.28) and guarantees the absolute convergence of the product (2.8).

*Step 5.* To show that the condition (2.9) is also sufficient for existence of the constant  $K$  in (2.11), recall that the function  $\kappa_0(t)$  in (1.6) is bounded. Assuming now that a bound on the function  $\kappa_k(t)$  exists, a bound on the function  $\kappa_{k+1}(t)$  is computed as follows: from the second equation in (2.7), the expression for  $\Phi_{k+1}(t, \sigma)$  and its inverse may be written as

$$\begin{aligned}\Phi_{k+1}(t, \sigma) &= (\mathbf{I} - \mathbf{R}_k(t))^{-1} \Phi_k(t, \sigma) (\mathbf{I} - \mathbf{R}_k(\sigma)), \\ \Phi_{k+1}^{-1}(t, \sigma) &= (\mathbf{I} - \mathbf{R}_k(\sigma))^{-1} \Phi_k^{-1}(t, \sigma) (\mathbf{I} - \mathbf{R}_k(t)).\end{aligned}\quad (2.30)$$

The substitution of (2.30) into the definition (2.11), followed by the use of the inequality (2.17) and the monotonicity of  $\phi^{[k]}$ , yields the following:

$$\begin{aligned}\kappa_{k+1}(t) &= \max_{0 < \sigma \leq t} \|\Phi_{k+1}(t, \sigma)\| \|\Phi_{k+1}^{-1}(t, \sigma)\| \\ &\leq \max_{0 < \sigma \leq t} \left( \|(\mathbf{I} - \mathbf{R}_k(t))^{-1}\| \|\Phi_k(t, \sigma)\| \|(\mathbf{I} - \mathbf{R}_k(\sigma))\| \right) \\ &\quad \times \left( \|(\mathbf{I} - \mathbf{R}_k(\sigma))^{-1}\| \|\Phi_k^{-1}(t, \sigma)\| \|\mathbf{I} - \mathbf{R}_k(t)\| \right) \\ &\leq \kappa_k(t) \|(\mathbf{I} - \mathbf{R}_k(t))^{-1}\| \|\mathbf{I} - \mathbf{R}_k(t)\| \max_{0 < \sigma \leq t} \|(\mathbf{I} - \mathbf{R}_k(\sigma))^{-1}\| \|\mathbf{I} - \mathbf{R}_k(\sigma)\| \\ &\leq \kappa_k(t) \frac{1 + r_k(t)}{1 - r_k(t)} \max_{0 < \sigma \leq t} \frac{1 + r_k(\sigma)}{1 - r_k(\sigma)} \\ &\leq \kappa_k(t) \frac{1 + \phi^{[k]}(\omega(t))}{1 - \phi^{[k]}(\omega(t))} \max_{0 < \sigma \leq t} \frac{1 + \phi^{[k]}(\omega(\sigma))}{1 - \phi^{[k]}(\omega(\sigma))} \\ &\leq \kappa_k(t) \left( \frac{1 + \phi^{[k]}(\omega(t))}{1 - \phi^{[k]}(\omega(t))} \right)^2.\end{aligned}\quad (2.31)$$

The last inequality, when iterated back to  $k = 0$ , generates the product,

$$\kappa_{k+1}(t) \leq \kappa_0(t) \left( \prod_{i=0}^k \frac{1 + \phi^{[i]}(\omega(t))}{1 - \phi^{[i]}(\omega(t))} \right)^2. \quad (2.32)$$

If  $\omega(t) \in [0, \omega_*)$ , then each factor in this product is bounded and positive, and so is  $\kappa_{k+1}(t)$ . As  $k \rightarrow \infty$ , the convergence of (2.32) depends on the convergence of a simpler product, namely,

$$\prod_{i=0}^k \frac{1 + \phi^{[i]}(\omega)}{1 - \phi^{[i]}(\omega)} = \prod_{i=0}^k 1 + \frac{2\phi^{[i]}(\omega)}{1 - \phi^{[i]}(\omega)}. \quad (2.33)$$

By the same theorem in [1], the simpler product in (2.33) converges provided the series  $\sum_{i=0}^{\infty} \phi^{[i]}(\omega)(1 - \phi^{[i]}(\omega))^{-1}$  converges. This is precisely the same series appearing in (2.29) and it can be bounded by the same function  $f(c\omega)$ . Therefore, given any  $\omega < \omega_*$



the limit function,  $\kappa_\infty(\omega(t))$ , exists and the constant  $K$  may be taken as the limit of (2.32), majorized by the inequalities (2.26),

$$K = \left( \prod_{i=0}^{\infty} \frac{c + (c\omega)^{2^i}}{c - (c\omega)^{2^i}} \right)^2 \max_{0 < \sigma \leq t} \kappa_0(\sigma). \quad \square$$

The foregoing inequalities simplify the proof of the following theorem.

**Theorem 2.** *The iteration (2.7) converges quadratically if  $\omega \in [0, \omega_*)$ .*

**Proof.** It suffices to consider the ratio  $\|\Delta_{k+1}(t)\|/\|\Delta_k(t)\|$ . From the norm of the last equation in (2.7), and the substitution of the inequalities (2.26) and  $c\omega < 1$ , it follows that

$$\frac{\|\Delta_{k+1}(t)\|}{\|\Delta_k(t)\|} \leq \frac{\|\mathbf{R}_k(t)\|}{1 - \|\mathbf{R}_k(t)\|} \leq \frac{(c\omega)^{2^k}}{c - (c\omega)^{2^k}} \leq \frac{(c\omega)^{2^k}}{c - 1}. \quad \square$$

### 3. The modified relative error and its product representation

The modified relative error matrix is defined by  $\mathbf{R}(t) = \Phi(t)\Phi_0^{-1}(t) - \mathbf{I}$ , so that  $\Phi(t)$  can be written as

$$\Phi(t) = (\mathbf{I} + \mathbf{R}(t))\Phi_0(t).$$

The analysis of this case is very similar to the previous case, hence only the differences will be highlighted. The differential equation for  $\mathbf{R}(t)$  is now

$$\mathbf{R}' = \Delta_0 + [\mathbf{A}_0, \mathbf{R}] + \Delta_0 \mathbf{R}, \quad \mathbf{R}(0) = \mathbf{0}. \quad (3.1)$$

Although the integrable part of (3.1) is the same as in (2.4), the solution and its L-perturbation are now evaluated according to:

$$\begin{aligned} \mathbf{R}_k(t) &= \int_0^t \Phi_k(t, \sigma) \Delta_k(\sigma) \Phi_k^{-1}(t, \sigma) d\sigma, \\ \Phi_{k+1}(t) &= (\mathbf{I} + \mathbf{R}_k(t))\Phi_k(t), \\ \Delta_{k+1}(t) &= \Delta_k(t)\mathbf{R}_k(t)(\mathbf{I} + \mathbf{R}_k(t))^{-1}. \end{aligned} \quad (3.2)$$

The second equation in (3.2) generates another left ordered infinite product representation for the fundamental matrix of solutions,

$$\Phi(t) = \lim_{m \rightarrow \infty} \left( \prod_{k=0}^m \mathbf{I} + \mathbf{R}_k(t) \right) \Phi_0(t). \quad (3.3)$$

**Theorem 3.** *The infinite product (3.3) is absolutely convergent provided*

$$\omega(t) < \omega_* = 0.79681213002 \dots \quad (3.4)$$

This is the same condition as in (2.9) of Theorem 1, hence the five steps in the proof below remain virtually the same except for a few differences.

**Proof.** By the same theorem in [1], the absolute convergence of the series,

$$\sum_{k=0}^{\infty} \|\mathbf{R}_k(t)\| < \infty, \quad (3.5)$$

near the origin is a sufficient condition for the absolute convergence of the product (3.3). The first step of the proof is the same as in Theorem 1.

*Steps 2 and 3.* The initial bound,  $r_0(t) \leq \omega(t)$ , remains the same and so does the bound on  $\delta_1(t) \leq \omega(t)(1 - \omega(t))^{-1} \omega'(t)/K$  in (2.13). This is not accidental because the difference between  $\mathbf{\Delta}_1(t)$  in (2.7) and the same matrix in (3.2) is only in the sign and the order of factors, and these disappear during evaluation of the norm  $\delta_1(t)$ . It follows, therefore, that for each  $k \geq 0$ ,  $r_k(t) \leq \phi^{[k]}(\omega(t))$ , and the rest of the steps 2 and 3 remain the same as in the proof of Theorem 1.

*Step 4.* To prove the absolute convergence of the series (3.5), write

$$\sum_{k=0}^{\infty} \|\mathbf{R}_k(t)\| \leq \sum_{k=0}^{\infty} \mathbf{r}_k(t) \leq \omega(t) + \sum_{k=1}^{\infty} \phi^{[k]}(\omega(t)). \quad (3.6)$$

Since all the functional iterates exist and are nonnegative, provided  $\omega(t) \in [0, \omega_*]$ , the substitution of the upper bounds (2.26) into (3.6) is valid, and the result is

$$\sum_{k=0}^{\infty} \|\mathbf{R}_k(t)\| \leq \omega(t) + \frac{1}{c} \sum_{k=1}^{\infty} (c\omega(t))^{2^k} = \frac{1}{c} f(c\omega(t)). \quad (3.7)$$

Here, the analytic function

$$f(z) = z + z^2 + z^4 + z^8 + \cdots = \sum_{k=0}^{\infty} z^{2^k}, \quad |z| < 1,$$

generated by  $g(z) = z$ , has a natural boundary at  $|z| = 1$  and it is real, nonnegative, and monotonically increasing on the real interval  $[0, 1)$ . This implies that, as long as  $c\omega \in [0, 1)$ , or  $\omega < \omega_*$ , the series (3.7) is bounded by  $c^{-1} f(c\omega(t))$  and the product (3.3) remains absolutely convergent.

*Step 5.* This step is virtually the same as in the proof of Theorem 1.  $\square$

#### 4. The A-relative error and its product representation

The A-relative error is defined by  $\mathbf{R} = 2(\Phi - \Phi_0)(\Phi + \Phi_0)^{-1}$ , provided that the matrix  $\Phi + \Phi_0$  is invertible. Solving for  $\Phi$  yields

$$\Phi(t) = \left( \frac{2 + \mathbf{R}(t)}{2 - \mathbf{R}(t)} \right) \Phi_0(t). \quad (4.1)$$

The differential equation for  $\mathbf{R}(t)$  is now a matrix Riccati equation,

$$\mathbf{R}' = \mathbf{\Delta}_0 + [\mathbf{A}_0, \mathbf{R}] + \frac{1}{2}[\mathbf{\Delta}_0, \mathbf{R}] - \frac{1}{4}\mathbf{R}\mathbf{\Delta}_0\mathbf{R}, \quad \mathbf{R}(0) = \mathbf{0}. \quad (4.2)$$

Nevertheless, its integrable part,  $\mathbf{R}_0$ , is the same as in (2.4) and when substituted into (4.1) it generates the first correction to the initial approximation,

$$\Phi_1(t) = \left( \frac{2 + \mathbf{R}_0(t)}{2 - \mathbf{R}_0(t)} \right) \Phi_0(t). \quad (4.3)$$

A lengthy computation of the matrix  $\mathbf{\Delta}_1(t) = \mathbf{A}(t) - \mathbf{L}[\Phi_1(t)]$  shows [3] that there is a general linear relation between consecutive L-perturbations,

$$\mathbf{\Delta}_1(t) = -(2 - \mathbf{R}_0)^{-1} (2[\mathbf{R}_0, \mathbf{\Delta}_0] + \mathbf{R}_0\mathbf{\Delta}_0\mathbf{R}_0) (2 + \mathbf{R}_0)^{-1}. \quad (4.4)$$

Equations (2.4), (4.3) and (4.4) define the iterative process that refines any admissible initial approximation,  $\Phi_0(t)$ , and its L-perturbation,  $\mathbf{\Delta}_0(t)$ :

$$\begin{aligned} \mathbf{R}_k(t) &= \int_0^t \Phi_k(t, \sigma) \mathbf{\Delta}_k(\sigma) \Phi_k^{-1}(t, \sigma) d\sigma, \\ \Phi_{k+1}(t) &= \frac{2 + \mathbf{R}_k(t)}{2 - \mathbf{R}_k(t)} \Phi_k(t), \\ \mathbf{\Delta}_{k+1}(t) &= -(2 - \mathbf{R}_k)^{-1} (2[\mathbf{R}_k, \mathbf{\Delta}_k] + \mathbf{R}_k\mathbf{\Delta}_k\mathbf{R}_k) (2 + \mathbf{R}_k)^{-1}. \end{aligned} \quad (4.5)$$

The second equation in (4.5) generates the third left ordered infinite product representation,

$$\Phi(t) = \lim_{m \rightarrow \infty} \left( \prod_{k=0}^m \frac{2 + \mathbf{R}_k(t)}{2 - \mathbf{R}_k(t)} \right) \Phi_0(t). \quad (4.6)$$

**Theorem 4.** *The infinite product (4.6) is absolutely convergent provided*

$$\omega(t) < \omega_\diamond = 0.84633199171\dots \quad (4.7)$$

**Proof.** Write each factor of (4.6) as  $(2 + \mathbf{R}_k)/(2 - \mathbf{R}_k) = \mathbf{I} + 2\mathbf{R}_k/(2 - \mathbf{R}_k)$ . By [1], if the matrix series  $\sum_{k=0}^{\infty} \mathbf{R}_k(2 - \mathbf{R}_k)^{-1}$  converges absolutely, i.e.,

$$\sum_{k=0}^{\infty} \|\mathbf{R}_k(2 - \mathbf{R}_k)^{-1}\| \leq \sum_{k=0}^{\infty} \frac{\|\mathbf{R}_k\|}{2 - \|\mathbf{R}_k\|} < \infty, \quad (4.8)$$

then so will the product. The first step is the same as in Theorem 1.

*Step 2.* The bound on  $\|\mathbf{R}_0(t)\| = r_0(t) \leq \omega(t)$  is also the same, but the third equation (4.5) yields a different bound on  $\delta_1(t) = \|\mathbf{\Delta}_1(t)\|$ :

$$\begin{aligned} \delta_1(t) &\leq \|(2 - \mathbf{R}_0)^{-1}\| \|2[\mathbf{R}_0, \mathbf{\Delta}_0] + \mathbf{R}_0\mathbf{\Delta}_0\mathbf{R}_0\| \|(2 + \mathbf{R}_0)^{-1}\| \\ &\leq \frac{4r_0(t) + (r_0(t))^2}{(2 - r_0(t))^2} \delta t^p \leq \frac{4\omega(t) + \omega^2(t)}{(2 - \omega(t))^2} \frac{\omega'(t)}{K}. \end{aligned} \quad (4.9)$$

With this inequality, the bound on  $r_1(t) = \|\mathbf{R}_1(t)\|$  is evaluated as follows:

$$\begin{aligned}
r_1(t) &\leq \int_0^t \|\Phi_1(t, \sigma)\| \|\Delta_1(\sigma)\| \|\Phi_1^{-1}(t, \sigma)\| d\sigma \\
&\leq \int_0^t \frac{4\omega(\sigma) + \omega^2(\sigma)}{(2 - \omega(\sigma))^2} d(\omega(\sigma)) = \psi(\omega(t)) - \psi(\omega(0)).
\end{aligned} \tag{4.10}$$

Here, the last integral is evaluated exactly in terms of the analytic function,

$$\psi(\omega) = \omega + \frac{6\omega}{2 - \omega} + 8 \log\left(1 - \frac{\omega}{2}\right) = \frac{\omega^2}{2} + \frac{5}{12}\omega^3 + \frac{1}{4}\omega^4 + \dots \tag{4.11}$$

All the coefficients in the series (4.11) are positive, therefore the function  $\psi$  is positive and monotonically increasing for  $\omega \in (0, 2)$ . By (1.7),  $\psi(\omega(0)) = \psi(0) = 0$ , hence (4.10) may be written as

$$r_1(t) \leq \psi(\omega(t)) = \psi^{[1]}(\omega(t)). \tag{4.12}$$

Since  $r_0(t) \leq \omega(t) = \psi^{[0]}(\omega(t))$ , inequality (4.12) sets the stage for an induction to prove a general bound for any  $k > 0$ , namely

$$r_k(t) \leq \psi^{[k]}(\omega(t)), \quad \psi^{[k]}(\omega) = \psi(\psi(\dots \psi(\omega) \dots)). \tag{4.13}$$

From this point on, the proof proceeds in exactly the same manner as in Theorem 1, with the iteration function  $\psi$  instead of  $\phi$ . This means that there will be a fixed point iteration,

$$\psi(\omega_{k+1}) = \omega_k, \quad \omega_0 = 2, \tag{4.14}$$

associated with the equation,

$$\psi(\omega) = \omega. \tag{4.15}$$

The iteration (4.14) generates the monotonically decreasing sequence,  $\{\omega_k\}_0^\infty$ , converging to the unique positive solution  $\omega_\diamond = 0.84633199\dots$  of (4.15). In other words, the inequality (4.7) is a sufficient condition for the existence and nonnegativity of all the functional iterates  $\psi^{[k]}$ .

*Step 3.* Like  $\phi^{[k]}$ , the iterates  $\psi^{[k]}$  satisfy a similar set of inequalities,

$$\begin{aligned}
2\left(\frac{\omega}{2}\right)^{2^k} &\leq \psi^{[k]} \leq \frac{1}{b}(b\omega)^{2^k}, \quad k > 0, \\
0 &\leq \dots \leq \psi^{[k]}(\omega) \leq \dots \leq \psi^{[1]}(\omega) \leq \omega \leq \omega_\diamond < 1,
\end{aligned} \tag{4.16}$$

over the interval  $[0, \omega_\diamond]$ . The only difference is that the constant  $b$  in the upper bound, which is determined from the interpolatory condition  $\psi(\omega_\diamond) = b\omega_\diamond^2 = \omega_\diamond$ , has the value  $b = \omega_\diamond^{-1} \approx 1.18156942$ .

*Step 4.* To bound the second sum in (4.8), use the inequalities (4.13) and the upper bounds (4.16) to obtain the following:

$$\sum_{k=0}^{\infty} \frac{\|\mathbf{R}_k(t)\|}{2 - \|\mathbf{R}_k(t)\|} \leq \sum_{k=0}^{\infty} \frac{r_k(t)}{2 - r_k(t)} \leq \sum_{k=0}^{\infty} \frac{\psi^{[k]}(\omega(t))}{2 - \psi^{[k]}(\omega(t))}$$

$$\leq \frac{\omega(t)}{2 - \omega(t)} + \sum_{k=1}^{\infty} \frac{(b\omega(t))^{2^k}}{2b - (b\omega(t))^{2^k}} = f(b\omega(t)). \quad (4.17)$$

In this case, the analytic function

$$f(z) = \frac{z}{2b - z} + \frac{z^2}{2b - z^2} + \frac{z^4}{2b - z^4} + \cdots = \sum_{k=0}^{\infty} \frac{z^{2^k}}{2b - z^{2^k}}, \quad |z| < 1,$$

generated by  $g(z) = z/(2b - z)$ , has also a natural boundary at  $|z| = 1$ . Therefore, as long as  $\omega(t) \in [0, \omega_\diamond)$ , the function  $f(b\omega(t))$  is real, nonnegative and monotonically increasing. It provides a bound for the series (4.8) and ensures the absolute convergence of the product (4.6).

*Step 5.* The condition (4.7) is also sufficient for the computation of the constant  $K$  in (2.11). To obtain a bound on  $\kappa_{k+1}(t)$  in terms of  $\kappa_k(t)$ , the computation follows exactly the same route as in Section 2, differing only in some of the details. For example, from the second equation (4.5), the expression for  $\Phi_{k+1}(t, \sigma)$  in terms of  $\Phi_k(t, \sigma)$  is now much longer,

$$\Phi_{k+1}(t, \sigma) = (2 + \mathbf{R}_k(t))(2 - \mathbf{R}_k(t))^{-1} \Phi_k(t, \sigma) (2 - \mathbf{R}_k(\sigma))(2 + \mathbf{R}_k(\sigma))^{-1}.$$

We shall skip a lengthy majorization process that starts by the substitution of the expressions for  $\Phi_{k+1}(t, \sigma)$  and  $\Phi_{k+1}^{-1}(t, \sigma)$  into the definition (2.11) for the function  $\kappa_{k+1}(t)$ . The steps during the bounding process (2.31) are the same, namely majorizations of matrix products, separation of the quantities that do not depend on  $\sigma$ , maximizations over  $\sigma$ , and the use of the monotonicity of  $\psi^{[k]}$ . The final result is

$$\kappa_{k+1}(t) \leq \kappa_k(t) \left( \frac{2 + \psi^{[k]}(\omega(t))}{2 - \psi^{[k]}(\omega(t))} \right)^4. \quad (4.18)$$

Iterating the inequality (4.18) back to  $k = 0$  generates the product,

$$\kappa_{k+1}(t) \leq \kappa_0(t) \left( \prod_{i=0}^k \frac{2 + \psi^{[i]}(\omega(t))}{2 - \psi^{[i]}(\omega(t))} \right)^4. \quad (4.19)$$

All the factors in the finite product 4.19 are positive and bounded if  $\omega \in [0, \omega_\diamond]$ . By the same arguments used in (2.32), and (2.33), as  $k \rightarrow \infty$ , the absolute convergence of the product (4.19) is assured provided the series  $\sum_{i=0}^{\infty} \psi^{[i]}(\omega(t))(2 - \psi^{[i]}(\omega(t)))^{-1}$  converges absolutely. This is the same series as in (4.17), hence it is bounded by the same function  $f(b\omega(t))$ . Consequently, for any  $\omega(t) < \omega_\diamond$ , the limit function  $\kappa_\infty(\omega(t))$  exists and the constant  $K$  may be computed by using the bounds (4.16) in the limit of (4.19):

$$K = \left( \prod_{i=0}^{\infty} \frac{2b + (b\omega)^{2^i}}{2b - (b\omega)^{2^i}} \right)^4 \max_{0 < \sigma \leq t} \kappa_0(\sigma), \quad \omega < \omega_\diamond. \quad \square$$

The rate of convergence of iteration (4.5) is quadratic provided  $\omega \in [0, \omega_\diamond)$ . The proof follows the same argument as in Theorem 2.

## 5. Conclusion

The ratio  $\omega_\diamond/\omega_* \approx 1.1082$  is not as impressive as one might have expected given the ability of the A-relative error to preserve properties for the unitary and symplectic systems, as well as having a larger denominator in the key equations (4.5) governing the iteration. It appears that a greater number of matrix inequalities employed in the case of the A-relative error, compared to the classical relative error, lessened the effect of its other advantages. In particular, the inequality  $\|[\mathbf{R}_k, \mathbf{\Delta}_k]\| \leq 2\|\mathbf{R}_k\|\|\mathbf{\Delta}_k\|$  deserves attention. Suppose this inequality is written as  $\|[\mathbf{R}_k, \mathbf{\Delta}_k]\| \leq 2s\|\mathbf{R}_k\|\|\mathbf{\Delta}_k\|$  so that  $s \in [0, 1]$  represents a scalar measure of noncommutativity, from the commutative case,  $s = 0$ , to the worst case,  $s = 1$ . Now, use this inequality to recalculate the function  $\psi$  in (4.11) as

$$\psi(\omega) = \omega + \frac{2\omega}{2-\omega} + 4\log\left(1 - \frac{\omega}{2}\right) + s\left(\frac{4\omega}{2-\omega} + 4\log\left(1 - \frac{\omega}{2}\right)\right).$$

The expansion of  $\psi$ , which splits into two series,

$$\psi(\omega) = s\left(\frac{\omega^2}{2} + \frac{\omega^3}{3} + \cdots\right) + \left(\frac{\omega^3}{12} + \cdots\right) = s\frac{\omega^2}{2} + O(\omega^3), \quad (5.1)$$

reveals that the presence of noncommutativity,  $s > 0$ , affects the quadratic term which is also the leading term in the asymptotic behavior of  $\psi$ . Table 1 shows the effect of  $s$  on the size of the fixed point  $\omega_\diamond(s)$ , which are computed from the equation  $\psi(\omega, s) = \omega$ , for a small set of uniformly spaced values of  $s \in [0, 1]$ .

Table 1

$s$	0.0	0.2	0.4	0.6	0.8	1.0
$\omega_\diamond(s)$	1.430664	1.261291	1.125276	1.014556	0.923050	0.846332

Finally, the absolute convergence of each product results in the inequality  $\omega(t) < \beta$ , where  $\beta$  is either  $\omega_*$  or  $\omega_\diamond$ . This can be inverted, by means of the definition (1.7), to obtain the explicit bound on the time interval for the absolute convergence,

$$t < \left(\frac{p+1}{\delta} \frac{\beta}{K}\right)^{\frac{1}{p+1}}.$$

The simplicity of this relation is entirely due to the assumed asymptotic behavior of the initial L-perturbation,  $\|\mathbf{\Delta}_0(t)\|$ , in (1.5). It should be noted that an asymptotic behavior, other than the polynomial growth in (1.5), would also yield computable bounds provided it is described by a sufficiently simple integrable function, but the calculation may be more involved.

## Acknowledgment

This work was partially supported by NASA under grant NGT5-40110 and cooperative agreement NCC5-581.

**References**

- [1] P. Henrici, Applied and Computational Complex Analysis, vol. 2, Wiley–Interscience, 1977.
- [2] Z.A. Melzak, Companion to Concrete Mathematics, Wiley, 1973.
- [3] I. Najfeld, W. Lakin, L-Derivative for an approximate solution to  $\Phi' = \mathbf{A}(t)\Phi$ : series and product formulae for left corrections, Quart. Appl. Math. LXI (2003) 537–564.